

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Exploring reasons for differences in performance between UK and International Medical Graduates in the Membership of the Royal College of General Practitioners Applied Knowledge Test: a cognitive interview study
AUTHORS	Pattinson, Julie; Blow, Carol; Sinha, Bijoy; Siriwardena, Aloysius

VERSION 1 - REVIEW

REVIEWER	Melvyn M Jones UCL Medical School UK
REVIEW RETURNED	15-Mar-2019

GENERAL COMMENTS	<p>This is a study exploring differences in approach to licensing exams between international medical graduates and home (UK) graduates. It is GP specific and may not be generalisable to other specialties. They use a qualitative approach using cognitive interviews, a "think aloud" approach to explore how candidates approached sample MRCGP AKT questions.</p> <p>Title does it need reference to GP/ MRCGP exams?</p> <p>The background seems comprehensive and captures much of the recent research in this area. The authors make a good case about what this study then adds.</p> <p>Methods these are relatively well described (see comments about Table 2 data). The response rate to invitations is not disclosed. The qualitative methods are robustly described. My main concern is whether you can compare 2 groups using quantitative methods? My understanding is this methodology would help explain the range of views or something of the "why" but not the "how much" questions, such as A different to B, which this would seem to be?</p> <p>Sampling 2/3rds of these GP trainees / registrars were ST1 who will probably not be working towards their AKT- most do it in ST2 or 3 years.</p> <p>The results are well described. Many of the issues identified would seem exam related and not specific to IMGs eg "recency"- this concept or difficulty would seem to be an age/ length of training issue not an IMG issue-but are IMGs older? The authors identify statistics training and lack of competence in this for IMGs; my experience is all AKT candidates feel inadequately prepared to do this.</p> <p>Some areas are frankly alarming e.g. that doctors in UK GP training are unaware of guidelines or NICE. This would suggest these candidates are not ready for the UK exit exam. I just wonder if this an artifact of their sampling of relatively junior GPSTs? Any candidate who was preparing seriously for AKT would have been at least prompted to this area by question banks etc. If true,</p>
-------------------------	---

	<p>however it does raise important questions about gaps in UK GP training.</p> <p>The discussion is reasonable but I think a much more cautious approach to interpreting the findings are required. I have listed some methodological concerns and I think a more nuanced approach is required.</p> <p>Issues</p> <p>I would check the RCGP is happy with the level of disclosure of elements of their exam? I know they have previously declared they do not release their exam questions. I would try to remove reference to the specifics of the exam questions eg "rupture rate of death" p 15 line 11.</p> <p>Table 2 is not referenced in the methods and is inadequately explained (nothing in methods) i.e. what is a "facility rate", this needs explanation. I am unsure whether with this purposive sampling any quantitative measure is appropriate. This table may be used to justify one groups superior performance over another and cannot be used on that context- I think it probably should be dropped.</p> <p>Minor issues abbreviations e.g. CT p11 line 33 verbatim quote but spell out.</p> <p>"Minor specialty topics" eg... paediatrics- I think the phrasing needs some serious work here p111 line 3-5.</p> <p>Overall I think this is an interesting study and adds to our knowledge of this important issue to the NHS- ensuring progress of GP trainees through exit exams and into practice. I have some misgiving of the comparative element of the study design but if the limitations of this approach are clearly described then the study probably should be published.</p>
--	---

REVIEWER	Lara Zibarras City University, London. UK.
REVIEW RETURNED	28-Mar-2019

GENERAL COMMENTS	<p>Overall, good & needed study, qualitative methodology. I suggest some minor amendments...</p> <p>Abstract</p> <p>Objectives</p> <p>I actually don't think you are investigating "causes"...</p> <p>As you state in the research question it's about the thought processes.. and whether there are differences. These don't amount to "causes" per se (I'd want a quantitative piece using structural equation modelling with a longitudinal study for a "cause" to be established). Please reword this section...</p> <p>Outcome measures - please reword as it lacks clarity. "...together with the problems or facilitators while doing so." this is vague and ambiguous. Please re-phrase.</p> <p>Method</p> <p>Please include a bit more about the "think aloud" questions, can you describe this in a bit more detail and perhaps include some of the questions you asked the participants.</p>
-------------------------	---

	<p>In the Settings & Participants section - did you tell participants that their results were anonymous? (you say it later, but don't specify here)</p> <p>Analysis - did you consider any researcher bias? I don't know enough about you as a research team, but did you consider this, or could this have been an issue in analysis? Did you do anything to minimise this?</p> <p>Results - Table 1 In the text on Page 10, I don't think Table 2 is explained well enough. Please include more clarity on rate of correct answers and difference in facility. Perhaps describing the positive and negative numbers too.</p> <p>Discussion - I feel you could have made more of your practical implications. You've only written a few lines here! Please build on what you think the practical implications are - like, should IMGs get extra help, should people struggling get help earlier, should they have cultural training? Practically speaking, how can we support people so that there aren't these differential rates. Or do you think these differential rates will just be there forever, and if so, what are the implications of that?</p> <p>Limitations section - also I think you should include that there was a small N, a very specific sample so it may not be generalisable, even to GP as a whole (with a different sample) and possibly not even to other healthcare settings. Would you recommend more research is done here? If so, what contexts, more using the same method?</p> <p>Finally, this paper might be helpful as well, although it doesn't directly discuss the AKT, there might be some important considerations. Patterson, F., Tiffin, P. A., Lopes, S., & Zibarras, L. (2018). Unpacking the dark variance of differential attainment on examinations in overseas graduates. Medical education, 52(7), 736-746.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1	
<p>This is a study exploring differences in approach to licensing exams between international medical graduates and home (UK) graduates. It is GP specific and may not be generalisable to other specialties. They use a qualitative approach using cognitive interviews, a “think aloud” approach to explore how candidates approached sample MRCGP AKT questions.</p>	<p>Thank you. We have added this to the Strengths and limitations, page 29, lines 562-564:</p> <p>Participants comprised a small sample of GPSTs with two thirds in their first year of training, so the results may not be generalisable to other specialties and may have been different for more experienced (second or third year) trainees.</p>
<p>The background seems comprehensive and captures much of the recent research in this area. The authors make a good case about what this study then adds.</p>	<p>Thank you.</p>

<p>Title does it need reference to GP/ MRCGP exams?</p>	<p>We have included the study design in the title and the title now has a reference to the MRCGP exam, page 1, lines 3-5:</p> <p>Exploring reasons for differences in performance between UK and International Medical Graduates in the Membership of the Royal College of General Practitioners; Applied Knowledge Test: a cognitive interview study</p>
<p>Methods: these are relatively well described.</p>	<p>Thank you.</p>
<p>The response rate to invitations is not disclosed.</p>	<p>In response we have included, page 10, line 164-166:</p> <p>The study was introduced to GPSTs during their weekly vocational training half-day educational programme and all those who volunteered to participate were offered an interview.</p>
<p>The qualitative methods are robustly described.</p>	<p>Thank you.</p>
<p>Sampling 2/3rds of these GP trainees / registrars were ST1 who will probably not be working towards their AKT- most do it in ST2 or 3 years.</p>	<p>We included on page 12-13, lines 223-227:</p> <p>We interviewed 21 GP specialty trainees (GPSTs: 8 female, 13 male), aged from 24 to 64 years, with two thirds in their first year of speciality training and the other third in years two or three. Of these, 13 participants were IMGs and 8 UK graduates. All IMGs and one UK-trained doctor were from a BME group (Table 1).</p> <p>Also in the Strengths and limitations, page 29, lines 562-564:</p> <p>Participants comprised a small sample of GPSTs with two thirds in their first year of training, so the results may not be generalisable to other specialities and may have been different for more experienced (second or third year) trainees.</p>
<p>Many of the issues identified would seem exam related and not specific to IMGs e.g. "recency"- this concept or difficulty would seem to be an age/ length of training issue not an IMG issue-but are IMGs older? The authors identify statistics training and lack of competence in this for IMGs; my experience is all AKT candidates feel inadequately prepared to do this.</p>	<p>We agree, and had emphasised this point in the discussion, page 26, lines 493-497:</p> <p>The key themes of real-life clinical experience, familiarity, and insight applied to both UK and IMG participants, while IMG participants experienced additional difficulties linked to differences in previous educational experience or familiarity with the UK NHS.</p>
<p>Some areas are frankly alarming e.g. that doctors in UK GP training are unaware of guidelines or NICE. This would suggest these candidates are not ready for the UK</p>	<p>We agree, although even the first year GP trainees will have had at least 7 years of medical training if they trained in the UK or possibly less (and certainly</p>

exit exam. I just wonder if this an artifact of their sampling of relatively junior GPSTs?	less exposure to UK guidelines) if they were IMGs having trained overseas.
Any candidate who was preparing seriously for AKT would have been at least prompted to this area by question banks etc. If true, however it does raise important questions about gaps in UK GP training.	We identified reasons why there may be gaps in UK training for IMGs and have expanded this in page 28, line 536-540: IMG doctors may lack preparation for entering UK speciality training for general practice, compared with UKGs who have previously had five years undergraduate medical training and an additional two years foundation training in the UK, for example in relation to learning about guidelines or adapting from a disease-centred to patient-centred model of care ^{22 39} .
The discussion is reasonable but I think a much more cautious approach to interpreting the findings are required. I have listed some methodological concerns and I think a more nuanced approach is required.	We agree and have revised the discussion and limitations in line with the recommendations and suggestions above and below.
Issues: I would check the RCGP is happy with the level of disclosure of elements of their exam? I know they have previously declared they do not release their exam questions. I would try to remove reference to the specifics of the exam questions e.g. "rupture rate of death" p 15 line 11.	The level of disclosure was agreed and has been approved by the MRCGP AKT clinical lead but we have removed this particular reference to "rupture rate of death" as suggested.
Table 2 is not referenced in the methods and is inadequately explained (nothing in methods) i.e. what is a "facility rate", this needs explanation. I am unsure whether with this purposive sampling any quantitative measure is appropriate. This table may be used to justify one group's superior performance over another and cannot be used on that context- I think it probably should be dropped.	Although we included this table to show how participants performance in the selected questions compared more generally with candidates' exam performance, it was not intended to justify one group's superior performance, and because of the potential for confusion we agree with the reviewer and have removed table 2 and relabelled the previous table 3 as table 2.
Minor issues abbreviations e.g. CT p11 line 33 verbatim quote but spell out.	We have spelt out the verbatim quote on page 14, line 263: "CT [Computed Tomography] in real life"
"Minor specialty topics" e.g.... paediatrics- I think the phrasing needs some serious work here p111 line 3-5.	We agree with the reviewer's comments and rephrased the sentence, page 14, line 247-251. Clinical exposure to specialties or speciality topics Limited clinical exposure to a host of speciality topics, including rheumatology, ophthalmology, paediatrics (managing medical conditions affecting infants, children and young people), gynaecology, relating to sexual health or radiology, provided difficulties answering questions for all participants.

<p>Reviewer: 2</p> <p>Overall, good & needed study, qualitative methodology.</p> <p>I suggest some minor amendments.</p>	<p>Thank you.</p>
<p>Abstract: Objectives</p> <p>I actually don't think you are investigating "causes"...</p> <p>As you state in the research question it's about the thought processes and whether there are differences. These don't amount to "causes" per se (I'd want a quantitative piece using structural equation modelling with a longitudinal study for a "cause" to be established). Please reword this section.</p>	<p>We agree that 'thought processes' is more appropriate and have reworded this section, page 3, lines 27-29:</p> <p>We aimed to investigate the thought processes of candidates answering multiple choice questions, considering differential attainment between IMGs and UK graduates.</p>
<p>Outcome measures - please reword as it lacks clarity. "...together with the problems or facilitators while doing so." this is vague and ambiguous. Please re-phrase.</p>	<p>We agree and have reworded this on page 4, line 41-42.</p> <p>Outcomes: Perceptions and views of participants on how they answered AKT questions together with any strategies used or difficulties experienced while doing so.</p>
<p>Method: Please include a bit more about the "think aloud" questions, can you describe this in a bit more detail and perhaps include some of the questions you asked the participants.</p>	<p>We included more detail and an example question, page 9, lines 137-142.</p> <p>The researcher followed an interview sequence where they asked a target question and used verbal probing to obtain more specific information (e.g. tell me a little bit more about why you think that is easy/difficult?) before moving forward to the next question. Target questions included, "Could you please talk me through in your own words how you perceive the standard introduction statement to the test and what it may (or may not) be telling you?" seeking comprehension of the question/complex instructions.</p>
<p>In the Settings & Participants section - did you tell participants that their results were anonymous? (you say it later, but don't specify here)</p>	<p>We have included in the Ethical considerations section, page 12, lines 214-216:</p> <p>Informed consent was taken from all participants and all participants were informed their data would be anonymised to prevent their identity being revealed.</p>
<p>Analysis - did you consider any researcher bias? I don't know enough about you as a research team, but did you consider this, or could this have been an issue in analysis? Did you do anything to minimise this?</p>	<p>We considered bias, and sought to minimise bias by ensuring that two researchers (JP/NS) coded the data, the research assistant JP was not a medical doctor, and did not have access to AKT answers, see page 12, line 206-2011:</p> <p>Initial in vivo coding,³³ where investigators used participants' individual wording and language to code a fragment of data, to reduce bias this was performed independently by two researchers (JP</p>

	and NS), but codes were subsequently compared to reach consensus. The interviewer (JP) also wrote reflective memos that helped with interpretation during data analysis. ²⁹ The interviewer (JP) was not a medical doctor and did not have access to the AKT answers until after the interviews were completed to minimise bias.
Results - In the text on Page 10, I don't think Table 2 is explained well enough. Please include more clarity on rate of correct answers and difference in facility. Perhaps describing the positive and negative numbers too.	Although we included table 2 to show how participants performance in the selected questions compared more generally with candidates' exam performance, because of the potential for confusion we have now removed table 2 and relabelled the previous table 3 as table 2.
Discussion - I feel you could have made more of your practical implications. You've only written a few lines here! Please build on what you think the practical implications are - like, should IMGs get extra help, should people struggling get help earlier, should they have cultural training? Practically speaking, how can we support people so that there aren't these differential rates? Or do you think these differential rates will just be there forever, and if so, what are the implications of that?	<p>We have expanded our discussion under Implications for future policy, research and practice, pages 28-29, line 566-576.</p> <p>This study provides information about the ways we can practically support all GP trainees including IMGs by highlighting gaps in training and experience and by identifying areas for intervention which may be helpful. The results also suggest wide differences in undergraduate experience which may disadvantage some doctors, particularly IMGs, for whom a standard three-year training programme may be insufficient or unrealistic to meet their needs. IMGs may require additional help prior to or early during GP training, to build cultural and interpersonal competence and confidence,⁴³ through familiarisation with NHS systems, clinical guidance, cultural or language differences and other areas where deficiencies in training, experience or learning approaches may leave them less prepared for licensing exams compared to UKGs. The costs of this early support could offset the additional costs of failure and extensions to training.</p>
Limitations section - also I think you should include that there was a small N, a very specific sample so it may not be generalisable, even to GP as a whole (with a different sample) and possibly not even to other healthcare settings. Would you recommend more research is done here? If so, what contexts, more using the same method?	<p>We have expanded the limitations as follows, page 29, lines 562-564:</p> <p>Participants comprised a small sample of GPSTs with two thirds in their first year of training, so the results may not be generalisable to other specialities and may have been different for more experienced (second or third year) trainees.</p>
Finally, this paper might be helpful as well, although it doesn't directly discuss the AKT, there might be some important considerations.	We agree and have referenced this paper as, although it does not directly discuss the AKT, it does acknowledge that IMGs require additional support in UK GP training, page 29, lines 571-572:

Patterson, F., Tiffin, P. A., Lopes, S., & Zibarras, L. (2018). Unpacking the dark variance of differential attainment on examinations in overseas graduates. Medical education, 52(7), 736-746.	IMGs may require additional help prior to or early during GP training, to build cultural and interpersonal competence and confidence, ⁴³ ...
--	---

VERSION 2 – REVIEW

REVIEWER	Melvyn Jones UCL Medical School UK
REVIEW RETURNED	23-Apr-2019

GENERAL COMMENTS	The authors have largely addressed my concerns. I still feel a % response rate needs to be included.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer comment:

The authors have largely addressed my concerns. I still feel a % response rate needs to be included.

We have expanded the first sentence of the results to include this information, page 11, line 219:

We interviewed 21 GP specialty trainees (GPSTs: 8 female, 13 male), aged from 24 to 64 years, with two thirds in their first year of speciality training and the other third in years two or three, who agreed to participate from a total cohort of 72 trainees (29%)